

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-181998

(43)Date of publication of application : 21.07.1995

(51)Int.Cl. G10L 9/14

(21)Application number : 05-327898 (71)Applicant : SANYO ELECTRIC CO LTD

(22)Date of filing : 24.12.1993 (72)Inventor : INOUE TATSUO
SUGISHITA SHOZO

(54) METHOD FOR COMPRESSING AND EXPANDING VOICE TIME-BASE

(57)Abstract:

PURPOSE: To develop a voice compression and expansion system by which a voice waveform small in displacement and distortion can be obtained even when the compression/expansion ratio is close to 1.

CONSTITUTION: When a voice signal is compressed, voice signal waveforms F and G whose lengths are ≥ 2 times as large a pitch cycles and integral multiples are cut from both the position P1 of the current pointer of the inputted voice signal waveform and the position P2 which is one pitch cycle precedent to the pointer, and a couple of cut voice signal waveforms are weighted with proper W1 and W2 and added to generate one signal waveform E by compressing the two voice signal waveforms, and, the current pointer P1 is moved on the inputted voice signal waveform

corresponding to compressibility to obtain a pointer P4, and similar operation is performed to obtain a compressed voice signal.

LEGAL STATUS [Date of request for examination] 22.04.1997

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 2951181

[Date of registration] 09.07.1999

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

*** NOTICES ***

**JPO and INPIT are not responsible for any
damages caused by the use of this translation.**

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] Are said twice [more than] pitch period, respectively with both the location of the present pointer of a sound signal wave inputted at the time of compression of a sound signal, and the location to which only 1 pitch period went to this pointer, and the sound signal wave of the die length of an integral multiple is started. Give suitable weight, add the sound signal wave of the started pair, and one signal wave form which compressed said two sound signal waves is generated. Next, the voice time-base-compact approach characterized by acquiring the sound signal compressed by making said current pointer move on said inputted sound signal wave according to compressibility, and performing same actuation.

[Claim 2] Are said twice [more than] pitch period, respectively with both the location of the present pointer of a sound signal wave inputted at the time of expanding of a sound signal, and the location in which only 1 pitch period was to this pointer, and the sound signal wave of the die length of an integral multiple is started. Give suitable weight, add the sound signal wave of the started pair, and one signal wave form which compounded said two sound signal waves is generated. Next, the voice time-axis expanding approach characterized by acquiring the sound signal elongated by making said current pointer move on said inputted sound signal wave according to the rate of expanding, and performing same actuation.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the audio time base compaction or the audio expanding approach which mainly encodes the digital sound signal used for the utterance speed-conversion approach, a voice IC recorder, etc. of the voice which controls the utterance rate in the case of processing voice in broadcast, a movie, and music.

[0002]

[Description of the Prior Art] A conventional digital sound signal is made into time base compaction or the approach of elongating. IEEE Trans.Acoust.Speech, Signal Processing, vol.ASSP-27, pp.121-133, and Apr.'79"Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals" -- TDHS (Tme-Domain Harmonic Scaling) indicated by D.Malah work and the collection of the flexible method in the time-axis of the voice by the correlation method, and "evaluation" Institute of Electronics and Communication Engineers spring lecture drafts -- pp.9-16 and 1986.5.21 Naotaka Morita -- PICOLA (Pointer Interval Control OverLap and Add) indicated by Itakura ***** is advocated.

[0003] It is die length N_c and [0004] from P2, respectively the time of only the pitch period T progressing to the present pointer location $P1$ and this, as shown in the wave form chart of drawing 7 (a) when setting to T the pitch period beforehand extracted using the pitch extractor to the voice wave in said TDHS and setting compressibility (or rate of expanding) to r .

[Equation 1]

$$N_C = rT / (1 - r)$$

[0005] It is die length N_c by starting ***** wave A and B, giving the weight (windowing function shown by the drawing destructive line F2) which goes conversely to 1 the weight (windowing function shown by the drawing destructive line F1) which goes to 1-0 linearly linearly from 0 at one wave B to Wave A, and adding both. Wave C is acquired. Next, it is a pointer on said input wave (drawing 7 (a)) [0006]

[Equation 2]

$$N_C + T = T / (1 - r)$$

[0007] It is the following die length N_c by carrying out ***** migration, starting two waves of the pitch period T like the above, giving weight to these and adding to them. A wave is acquired. It is the input wave of die-length N_c+T to the die length N_c by repeating this actuation henceforth. An output wave is made continuously. And this die length N_c The wave is what satisfied compressibility r to the input wave of die-length N_c+T .

[0008] It is die length N_s and [0009] from P4, respectively the time of on the other hand, only the pitch period T being overdue to the present pointer location $P1$ and this in expanding, as shown in the wave form chart of drawing 8 (a) and (b).

[Equation 3]

$$N_S = rT / (r - 1)$$

[0010] It is die length N_S by starting ***** wave A and B, giving the weight (windowing function shown by the drawing destructive line F4) which goes conversely to 0 the weight (windowing function shown by the drawing destructive line F3) which goes to 0-1 linearly linearly from 1 at one wave B to Wave A, and adding both. Wave C is acquired. Next, it is a pointer on said input wave (drawing 8 (a)) [0011]

[Equation 4]

$$N_S - T = T / (r - 1)$$

[0012] It is the following die length N_S by carrying out **** migration, starting two waves of the pitch period T like the above, attaching weight to these and adding to them. A wave is acquired. The output wave of the rate r of expanding is continuously acquired to the input wave of die-length $N_S \cdot T$ by repeating this actuation henceforth. And this die length N_S The wave is what satisfied the rate r of expanding to the input wave of die-length $N_S \cdot T$.

[0013] By the way, if this TDHS is applied to compression or expanding of a time-axis, compressibility or the rate r of expanding is the wave length N_c as mentioned above, when close to 1. N_S It became larger than the pitch period T too much, the wave-like gap by fluctuation of a pitch period arose within the voice wave to add, and there was a trouble of becoming the cause of distortion.

[0014]

[Problem(s) to be Solved by the Invention] When TDHS is applied to compression or expanding of a time-axis as mentioned above, compressibility (or rate of expanding) r this invention When close to 1, N_c N_S The wave-like gap by fluctuation of a pitch period arises within the voice wave added more greatly than T . In view of the conventional trouble which is said when it comes to the cause of distortion, even when compressibility (or rate of expanding) is close to 1, it aims at developing the speech compression and the expanding method with which the small voice wave of a gap or distortion is acquired.

[0015]

[Means for Solving the Problem] This invention starts the sound signal wave of the die length of the more than twice and the integral multiple of said pitch period, respectively from both the location of the present pointer of a sound signal wave inputted at the time of compression of a sound signal, and the location to which only 1 pitch period went to this pointer. Give suitable

weight, add the sound signal wave of the started pair, and one signal wave form which compressed said two sound signal waves is generated. Next, said current pointer is made to move on said inputted sound signal wave or the generated sound signal according to compressibility, and the sound signal compressed by performing same actuation is acquired.

[0016] Moreover, this invention is said twice [more than] pitch period, respectively with both the location of the present pointer of a sound signal wave inputted at the time of expanding of a sound signal, and the location in which only 1 pitch period was to this pointer, and starts the sound signal wave of the die length of an integral multiple. Give suitable weight, add the sound signal wave of the started pair, and one signal wave form which compounded said two sound signal waves is generated. Next, said current pointer is made to move on said inputted sound signal wave or the generated sound signal according to the rate of expanding, and the sound signal elongated by performing same actuation is acquired.

[0017]

[Function] The die length of a voice wave which is applied according to the above-mentioned configuration is concerned with compressibility or the rate of expanding, and the distortion of the sound signal acquired as the waves which were twice [more than] the pitch period that there is nothing, and were the die length of an integral multiple, and were always most alike since adjustment of compressibility or the rate of expanding was performed by the movement magnitude of a pointer were added and compressibility or the rate of expanding approached 1 becomes small.

[0018]

[Example] The voice time-base-compaction approach and the expanding approach of this invention are explained to a detail based on a drawing about the one example below.

The Fig. corresponding to the [compression approach] shows the circuit block diagram with which drawing 1 - drawing 3 drawing 1 realizes the compression approach concerned.

[0019] Analog sound signal $s(t)$ is changed into digitized voice data $s(n)$ by A/D converter 1 in drawing 1. Next, the pitch period T of digitized voice data $s(n)$ is extracted by the pitch extractor 2, time base compaction of voice data $s(n)$ is performed by the time-base-compaction machine 3 using the extracted pitch information, and the digitized voice signal $s_c(n)$ compressed as a result is acquired.

[0020] And said pitch period information T and the compressed sound signal $s_c(n)$ are changed into D/A converter 4 delivery and here at analog sound signal $s'(t)$, and is outputted.

[0021] Detailed work of the time-axis compressor 3 in said drawing 1 is explained using drawing

2 (a), (b), and drawing 3 (a) and (b) below.

[0022] As said time-axis compressor 3 shows to drawing 2 (a), the pointer P1 of the present respectively, on the other hand, the time of only the pitch period T progressing -- from P2 -- respectively -- the voice wave F of die-length 2T (the voice wave A+ voice wave B --) namely, the section of pointers P1-P3 and the voice wave G (the voice wave B+ voice wave C --) That is, the section of pointers P2-P4 is started, the weight (windowing function W1) which goes to 1-0 is given to Wave F, the weight (windowing function W2) which goes to 1 from 0 conversely is given to Wave G, both [these] waves are added, and the voice wave E of die-length 2T is acquired (refer to drawing 2 (b)). In addition, the weight W1 and W2 here is formed in order to maintain the wave-like continuity in the node before and behind the voice wave E.

[0023] Next, Lc of drawing 2 [0024]

[Equation 5]

$$L_C = rT / (1 - r)$$

[0025] the time of carrying out -- the time of $L_c \geq 2T$ -- the input voice top of said drawing 2 (a) -- $T+L_c$ only -- a pointer is moved and the actuation same as a pointer P5 as the above is performed. The output at this time will output said wave E and the input waves of die-length L_c-2T from P4 to P5 following this as they are.

[0026] the input voice top which shows drawing 3 (a) on the other hand at the time of $L_c < 2T$ -- a pointer P1 to $T+L_c$ only -- a pointer is moved to the right and the actuation same as a pointer P6 as drawing 2 (a) is performed. The output at this time is the die length Lc in the first half of Wave E. It becomes only a part and they are remaining $2T-L_c$. Since a part is again used for the next processing, it is returned to an input side.

[0027] It is die-length $T+L_c$ by the above actuation. An input wave to die length Lc The output wave was made and this has satisfied the compressibility r of a sound signal. In addition, compressibility $r=3/4$, drawing 3 (a), and (b) show the case of compressibility $r=3/5$, respectively, and, as for drawing 2 (a) and (b), the wave length which starts a compressional-wave form shows the case of the twice of a pitch period for both sides.

[0028] It cannot be overemphasized that it is not what is restricted to this example that the wave started, of course should just be a twice [more than] as many integral multiple as this.

The Fig. corresponding to the [expanding approach] shows the circuit block diagram with which drawing 4 · drawing 6 drawing 4 realizes the expanding approach concerned.

[0029] Analog sound signal s (t) is changed into digitized voice data s (n) by A/D converter 1 in

drawing 4 . Next, the pitch period T of digitized voice data s (n) is extracted by the pitch extractor 2, time-axis expanding of voice data s (n) is performed by the time-axis expander 5 using the extracted pitch information, and the digitized voice signal se (n) elongated as a result is acquired. [0030] And said pitch period information T and the elongated sound signal se (n) are changed into D/A converter 4 delivery and here at analog sound signal s' (t), and is outputted.

[0031] Detailed work of the time-axis expander 5 in said drawing 4 is explained using drawing 5 (a), (b), and drawing 6 (a) and (b) below.

[0032] As said time-axis expander 5 shows to drawing 5 (a), the present pointer P1, the time of only the pitch period T being overdue to this pointer P1 -- from P4 -- respectively -- the voice wave F of die-length 2T (the wave A+ wave B --) namely, the section of pointers P4-P2 and the voice wave G (the wave B+ wave C --) That is, the section of pointers P1-P3 is started, and the voice wave H of die-length 2T which gives the weight W4 which tends toward Wave G weight W3 which goes to 0.1 1.0 to Wave F, and is shown in drawing 5 (b) in addition is created. Weight W3 at this time and W4 are prepared in order to keep the continuity in the node before and behind the voice wave H being the same as that of the case of said time-axis compressor 3.

[0033] And Ls [0034]

[Equation 6]

$$L_s = T / (r - 1)$$

[0035] the time of carrying out -- the time of $L_s \geq T$ -- the input voice top of said drawing 5 (a) -- Ls only -- it is made to move and actuation same as a pointer P3 is performed. The output at this time outputs Wave H and the input wave (wave C) of die-length Ls-T to the pointers P2-P3 following this as it is. on the other hand -- the time of $L_s < T$ -- the input voice top of drawing 6 (a) -- a pointer P1 to Ls only -- a pointer is moved and the actuation same as a pointer P5 as the above is performed. The output at this time is die-length T+Ls in the first half of Wave H. It becomes only a part and is remaining T-Ls. Since a part is again used for the next processing, it is returned to an input side.

[0036] It is die length Ls by the above actuation. An input wave to die-length T+Ls The output wave is made and it turns out that the rate r of expanding is satisfied. In addition, drawing 5 (a) and (b) show the case of rate r=of expanding 3/2, and the wave length which starts when both drawing 6 (a) and (b) make an expanding wave from the case of the rate r= 3 of expanding shows the case of the twice of a pitch period. It cannot be overemphasized that it is not what is restricted to this example that the wave started, of course should just be a twice [more than] as

many integral multiple as this.

[0037]

[Effect of the Invention] The die length of a voice wave applied at the time of compression or expanding is concerned with compressibility or the rate of expanding, it is the die length of a twice [more than] as many integral multiple as a pitch period that there is nothing, and since compression or adjustment of expanding is performed by the movement magnitude of a pointer, the wave-like gap by fluctuation of a pitch period is suppressed, and, according to the time-base-compaction approach of this invention, and the expanding approach, it can expect the effectiveness that the distortion of a voice wave at the time of compression or expanding becomes small as a result.

[0038] Moreover, the effectiveness that distortion becomes small is also expectable as compressibility or the rate of expanding approaches 1 as compared with the conventional approach.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block circuit diagram which realizes the time-base-compaction approach of this invention.

[Drawing 2] It is drawing showing input wave (a) at the time of compression by this invention approach ($L_c \geq 2T$), and output wave (b).

[Drawing 3] It is drawing showing input wave (a) at the time of compression by this invention approach ($L_c < 2T$), and output wave (b).

[Drawing 4] It is the block circuit diagram which realizes the time-axis expanding approach of this invention.

[Drawing 5] It is drawing showing input wave (a) at the time of expanding by this invention approach ($L_s \geq T$), and output wave (b).

[Drawing 6] It is drawing showing input wave (a) at the time of expanding by this invention approach ($L_s < T$), and output wave (b).

[Drawing 7] It is drawing of the conventional approach corresponding to drawing 2 and drawing 3.

[Drawing 8] It is drawing of the conventional approach corresponding to drawing 5 and drawing 6.

[Description of Notations]

1 A/D Converter

2 Pitch Extractor

3 Time-axis Compressor

4 D/A Converter

5 Time-axis Expander

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-181998

(43) 公開日 平成7年(1995)7月21日

(51) Int.Cl.⁶

識別記号

序内整理番号

F I

技術表示箇所

G 1 0 L 9/14

G
J

審査請求 未請求 請求項の数 2 O L (全 7 頁)

(21) 出願番号 特願平5-327898

(22) 出願日 平成5年(1993)12月24日

(71) 出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(72) 発明者 井上 健生

大阪府守口市京阪本通2丁目18番地 三洋
電機株式会社内

(72) 発明者 杉下 正蔵

大阪府守口市京阪本通2丁目18番地 三洋
電機株式会社内

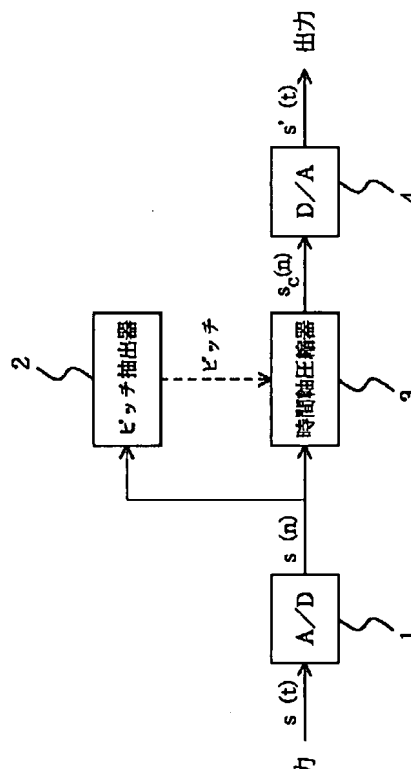
(74) 代理人 弁理士 岡田 敬

(54) 【発明の名称】 音声時間軸圧縮方法及び伸長方法

(57) 【要約】

【目的】 圧縮伸長率が1に近い時でもずれや歪みの小さい音声波形が得られる音声圧縮伸長方式を開発する。

【構成】 音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置P1と該ポイントに対して1ピッチ周期だけ進んだ位置P2との両方から夫々前記ピッチ周期の2倍以上且つ整数倍の長さの音声信号波形F、Gを切り出し、切り出された一対の音声信号波形を適当な重みW1、W2をつけて加算し、前記2つの音声信号波形を圧縮した1つの信号波形Eを生成し、次に前記現在ポイントP1を圧縮率に応じて前記入力された音声信号波形上で移動せしめてポイントP4として同様の操作を行い圧縮された音声信号を得る。



【特許請求の範囲】

【請求項 1】 音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して 1 ピッチ周期だけ進んだ位置との両方から夫々前記ピッチ周期の 2 倍以上で且つ整数倍の長さの音声信号波形を切り出し、切り出された一对の音声信号波形を適当な重みをつけて加算し、前記 2 つの音声信号波形を圧縮した 1 つの信号波形を生成し、次に前記現在ポイントを圧縮率に応じて前記入力された音声信号波形上で移動せしめて同様の操作を行い圧縮された音声信号を得ることを特徴とする音声時間軸圧縮方法。

【請求項 2】 音声信号の伸長時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して 1 ピッチ周期だけ遅れた位置との両方から夫々前記ピッチ周期の 2 倍以上で且つ整数倍の長さの音声信号波形を切り出し、切り出された一对の音声信号波形を適当な重みをつけて加算し、前記 2 つの音声信号波形を合成した 1 つの信号波形を生成し、次に前記現在ポイントを伸長率に応じて前記入力された音声信号波形上で移動せしめて同様の操作を行い伸長された音声信号を得ることを特徴とする音声時間軸伸長方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、放送、映画、音楽において、音声処理する場合の発声速度を制御する音声の発声速度変換方法、及び音声 IC レコーダ等に用いられるデジタルの音声信号を主として符号化する音声の時間軸圧縮あるいは伸長方法に関する。

【0002】

【従来の技術】従来デジタルの音声信号を時間軸圧縮あるいは伸長する方法として、IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-27, pp. 121-133, Apr. '79 "Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals" D. Malah 著に開示された TDHS (Time-Domain Harmonic Scaling) や「自己相関法による音声の時間軸での伸縮方式とその評価」電子通信学会春季講演予稿集 pp. 9-16, 1986. 5.21 森田直孝、板倉文忠著に開示された PICOLA (Pointer Interval Control Overlap and Add) 等が提唱されている。

【0003】前記 TDHS では、音声波形に対して予めピッチ抽出器を用いて抽出されたピッチ周期を T 、圧縮率（あるいは伸長率）を r とするとき、図 7 (a) の波形図に示すように現在のポイント位置 P_1 と、これに対してピッチ周期 T だけ進んだ時点 P_2 から夫々長さ N_c 、

【0004】

【数 1】

$$N_c = rT / (1 - r)$$

【0005】の音声波形 A 及び B を切り出し、波形 A には 1 から 0 へ直線的に向かう重み（図中破線 F 1 で示す窓関数）を、一方の波形 B には逆に 0 から 1 へ直線的に向かう重み（図中破線 F 2 で示す窓関数）を付与して、両者を加え合わせるにより長さ N_c の波形 C を得る。次にポイントを前記入力波形（図 7 (a)）上で、

【0006】

【数 2】

$$N_c + T = T / (1 - r)$$

【0007】だけ移動して、前記と同様にピッチ周期 T の 2 つの波形を切り出し、これらに重みをつけて加え合わせるにより、次の長さ N_c の波形を得る。以後この操作を繰り返すことにより長さ $N_c + T$ の入力波形から長さ N_c の出力波形が連続して作り出される。しかもこの長さ N_c の波形は長さ $N_c + T$ の入力波形に対して圧縮率 r を満足したものとなっている。

【0008】一方、伸長の場合には図 8 (a) (b) の波形図に示すように現在のポイント位置 P_1 とこれに対してピッチ周期 T だけ遅れた時点 P_4 から夫々長さ N_s 、

【0009】

【数 3】

$$N_s = rT / (r - 1)$$

【0010】の音声波形 A 及び B を切り出し、波形 A には 0 から 1 へ直線的に向かう重み（図中破線 F 3 で示す窓関数）を、一方の波形 B には逆に 1 から 0 へ直線的に向かう重み（図中破線 F 4 で示す窓関数）を付与して、両者を加え合わせるにより長さ N_s の波形 C を得る。次にポイントを前記入力波形（図 8 (a)）上で、

【0011】

【数 4】

$$N_s - T = T / (r - 1)$$

【0012】だけ移動して、前記と同様にピッチ周期 T の 2 つの波形を切り出し、これらに重みを付けて加え合わせるにより、次の長さ N_s の波形を得る。以後この操作を繰り返すことにより長さ $N_s - T$ の入力波形に対し伸長率 r の出力波形が連続して得られる。しかもこの長さ N_s の波形は長さ $N_s - T$ の入力波形に対して伸長率 r を満足したものとなっている。

【0013】ところでかかる TDHS を時間軸の圧縮あるいは伸長に適用すると、上述のように圧縮率あるいは伸長率 r が 1 に近い場合、波形長 N_c や N_s がピッチ周期 T よりも大きくなり過ぎてしまい、加え合わせる音声波形内でピッチ周期の揺らぎによる波形のずれが生じ、歪みの原因になるという問題点があった。

【0014】

【発明の目的】本発明は、上述の問題点を解決し、歪みの原因にならないという問題点があった。

にTDHSを時間軸の圧縮あるいは伸長に適用すると圧縮率（あるいは伸長率） r が1に近い場合、 N_c や N_s が T より大きく加え合わせる音声波形内でピッチ周期の揺らぎによる波形のずれが生じ、歪みの原因となるという従来の問題点に鑑み、圧縮率（あるいは伸長率）が1に近い時でもずれや歪みの小さい音声波形が得られる音声圧縮及び伸長方式を開発することを目的とする。

【0015】

【課題を解決するための手段】本発明は、音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して1ピッチ周期だけ進んだ位置との両方から夫々前記ピッチ周期の2倍以上且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記2つの音声信号波形を圧縮した1つの信号波形を生成し、次に前記現在ポイントを圧縮率に応じて前記入力された音声信号波形上あるいは生成された音声信号上で移動せしめて同様の操作を行い圧縮された音声信号を得るものである。

【0016】また本発明は、音声信号の伸長時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して1ピッチ周期だけ遅れた位置との両方から夫々前記ピッチ周期の2倍以上で且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記2つの音声信号波形を合成した1つの信号波形を生成し、次に前記現在ポイントを伸長率に応じて前記入力された音声信号波形上あるいは生成された音声信号上で移動せしめて同様の操作を行い伸長された音声信号を得るものである。

【0017】

【作用】上記構成によると、加え合わせる音声波形の長さは、圧縮率あるいは伸長率に関わり無くピッチ周期の2倍以上で且つ整数倍の長さであり、ポイントの移動量によって圧縮率あるいは伸長率の調整が行われるため、常に最も似通った波形同士が加え合わされ、圧縮率あるいは伸長率が1に近づくに従って得られた音声信号の歪みは小さくなる。

【0018】

【実施例】以下本発明の音声時間軸圧縮方法及び伸長方法をその一実施例について図面に基づいて詳細に説明する。

〔圧縮方法〕対応図は図1～図3

図1は当該圧縮方法を実現する回路ブロック図を示す。

【0019】図1においてアナログ音声信号 $s(t)$ はA/D変換器1によりデジタル音声データ $s(n)$ に変換される。次にピッチ抽出器2によりデジタル音声データ $s(n)$ のピッチ周期 T が抽出され、抽出されたピッチ情報を用いて時間軸圧縮器3により音声データ $s(n)$ の時間軸圧縮が行われ、結果として圧縮されたデ

ジタル音声信号 $s_c(n)$ を得る。

【0020】そして前記ピッチ周期情報 T と圧縮された音声信号 $s_c(n)$ をD/A変換器4に送り、ここでアナログ音声信号 $s'(t)$ に変換して出力する。

【0021】以下に前記図1における時間軸圧縮器3の詳細な働きを図2(a)(b)及び図3(a)(b)を用いて説明する。

【0022】前記時間軸圧縮器3では図2(a)に示すように、夫々現在のポイント P_1 と、これに対してピッチ周期 T だけ進んだ時点 P_2 から、夫々長さ $2T$ の音声波形 F （音声波形A+音声波形B、即ちポイント $P_1 \sim P_3$ の区間）及び音声波形 G （音声波形B+音声波形C、即ちポイント $P_2 \sim P_4$ の区間）を切り出し、波形 F には1から0へ向かう重み（窓関数 W_1 ）をつけ、波形 G には逆に0から1へ向かう重み（窓関数 W_2 ）をつけてこれら両波形を加え合わせ、長さ $2T$ の音声波形 E を得る（図2(b)参照）。なお、ここでの重み W_1 、 W_2 は音声波形 E の前後の接続点での波形の連続性を保つために設けたものである。

【0023】次に図2の L_c を、

【0024】

【数5】

$$L_c = rT / (1 - r)$$

【0025】とすると、 $L_c \geq 2T$ のときには前記図2(a)の入力音声上で $T + L_c$ だけポイントを移動させてポイント P_5 として上記と同様の操作を行う。この時の出力は前記波形 E とこれに続く P_4 から P_5 までの長さ $L_c - 2T$ の入力波形をそのまま出力することになる。

【0026】一方 $L_c < 2T$ のときは図3(a)に示す入力音声上でポイント P_1 から $T + L_c$ だけ右にポイントを移動させてポイント P_6 として図2(a)と同様の操作を行う。この時の出力は波形 E の前半の長さ L_c の部分のみとなり、残りの $2T - L_c$ の部分は次の処理に再び用いられるため入力側へ戻される。

【0027】以上の操作によって長さ $T + L_c$ の入力波形から長さ L_c の出力波形が作られ、これは音声信号の圧縮率 r を満足するものとなっている。なお、図2

(a)(b)は圧縮率 $r = 3/4$ 、図3(a)(b)は圧縮率 $r = 3/5$ の場合を夫々示しており、双方とも圧縮波形を切り出す波形長はピッチ周期の2倍の場合を示している。

【0028】もちろん切り出す波形は2倍以上の整数倍であれば良くこの実施例に限られるものではないことは言うまでもない。

〔伸長方法〕対応図は図4～図6

図4は当該伸長方法を実現する回路ブロック図を示す。

【0029】図4においてアナログ音声信号 $s(t)$ はA/D変換器1によりデジタル音声データ $s(n)$ に変換される。次にピッチ抽出器2によりデジタル音声データ $s(n)$ のピッチ周期 T が抽出され、抽出されたピッチ情報を用いて時間軸伸長器3により音声データ $s(n)$ の時間軸伸長が行われ、結果として伸長されたデ

ジタル音声信号 $s_s(n)$ を得る。

タ $s(n)$ のピッチ周期 T が抽出され、抽出されたピッチ情報を用いて時間軸伸長器 5 により音声データ s

(n) の時間軸伸長が行われ、結果として伸長されたデジタル音声信号 $s_e(n)$ を得る。

【0030】そして前記ピッチ周期情報 T と伸長された音声信号 $s_e(n)$ を D/A 変換器 4 に送り、ここでアナログ音声信号 $s'(t)$ に変換して出力する。

【0031】以下に前記図 4 における時間軸伸長器 5 の詳細な働きを図 5 (a) (b) 及び図 6 (a) (b) を用いて説明する。

【0032】前記時間軸伸長器 5 では図 5 (a) に示すように、現在のポインタ $P1$ と、このポインタ $P1$ に対してピッチ周期 T だけ遅れた時点 $P4$ から夫々長さ $2T$ の音声波形 F (波形 A + 波形 B 、即ちポインタ $P4 \sim P2$ の区間) 及び音声波形 G (波形 B + 波形 C 、即ちポインタ $P1 \sim P3$ の区間) を切り出し、波形 F には 0 から 1 へ向かう重み $W3$ を、波形 G には 1 から 0 へ向かう重み $W4$ をつけて加え合わせ、図 5 (b) に示す長さ $2T$ の音声波形 H を作成する。この時の重み $W3$ 、 $W4$ は前記時間軸圧縮器 3 の場合と同様に音声波形 H の前後の接続点での連続性を保つために設けるものである。

【0033】そして L_s を

【0034】

【数 6】

$$L_s = T / (r - 1)$$

【0035】とすると、 $L_s \geq T$ の時は前記図 5

(a) の入力音声上で L_s だけ移動させてポインタ $P3$ として同様の操作を行う。この時の出力は波形 H とこれに続くポインタ $P2 \sim P3$ までの長さ $L_s - T$ の入力波形 (波形 C) をそのまま出力する。一方、 $L_s < T$ のときは図 6 (a) の入力音声上でポインタ $P1$ から L_s だけポインタを移動させて、ポインタ $P5$ として上記と同様の操作を行う。この時の出力は、波形 H の前半の長さ $T + L_s$ の部分のみとなり、残りの $T - L_s$ の部分は次の処理に再び用いられるため、入力側に戻される。

【0036】以上の操作によって長さ L_s の入力波形から長さ $T + L_s$ の出力波形が作られており、伸長率 r が満足されていることが解る。なお、図 5 (a) (b) は

伸長率 $r = 3/2$ の場合を示しており、図 6 (a)

(b) は伸長率 $r = 3$ の場合で共に伸長波形を作る時に切り出す波形長はピッチ周期の 2 倍の場合を示している。もちろん切り出す波形は 2 倍以上の整数倍であれば良くこの実施例に限られるものではないことは言うまでもない。

【0037】

【発明の効果】本発明の時間軸圧縮方法及び伸長方法によると、圧縮時または伸長時に加え合わせる音声波形の長さは圧縮率あるいは伸長率に関わり無くピッチ周期の 2 倍以上の整数倍の長さであり、ポインタの移動量によって圧縮あるいは伸長の調整が行われるため、ピッチ周期の揺らぎによる波形のずれが抑えられ、結果的に圧縮あるいは伸長時の音声波形の歪みが小さくなる効果が期待できる。

【0038】また従来の方と比較して圧縮率あるいは伸長率が 1 に近づくに従って歪みが小さくなる効果も期待できる。

【図面の簡単な説明】

【図 1】本発明の時間軸圧縮方法を実現するブロック回路図である。

【図 2】本発明方法による圧縮時 ($L_c \geq 2T$) の入力波形 (a) 及び出力波形 (b) を示す図である。

【図 3】本発明方法による圧縮時 ($L_c < 2T$) の入力波形 (a) 及び出力波形 (b) を示す図である。

【図 4】本発明の時間軸伸長方法を実現するブロック回路図である。

【図 5】本発明方法による伸長時 ($L_s \geq T$) の入力波形 (a) 及び出力波形 (b) を示す図である。

【図 6】本発明方法による伸長時 ($L_s < T$) の入力波形 (a) 及び出力波形 (b) を示す図である。

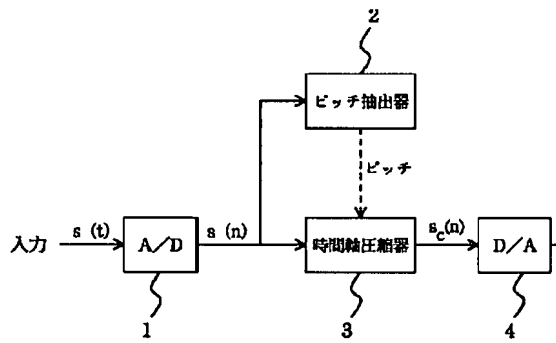
【図 7】図 2、図 3 に対応する従来方法の図である。

【図 8】図 5、図 6 に対応する従来方法の図である。

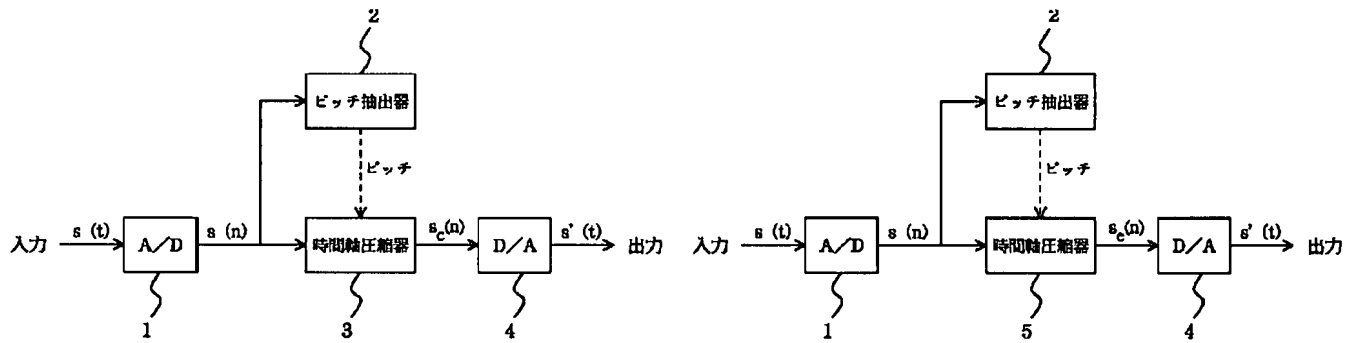
【符号の説明】

1	A/D 変換器
2	ピッチ抽出器
3	時間軸圧縮器
4	D/A 変換器
5	時間軸伸長器

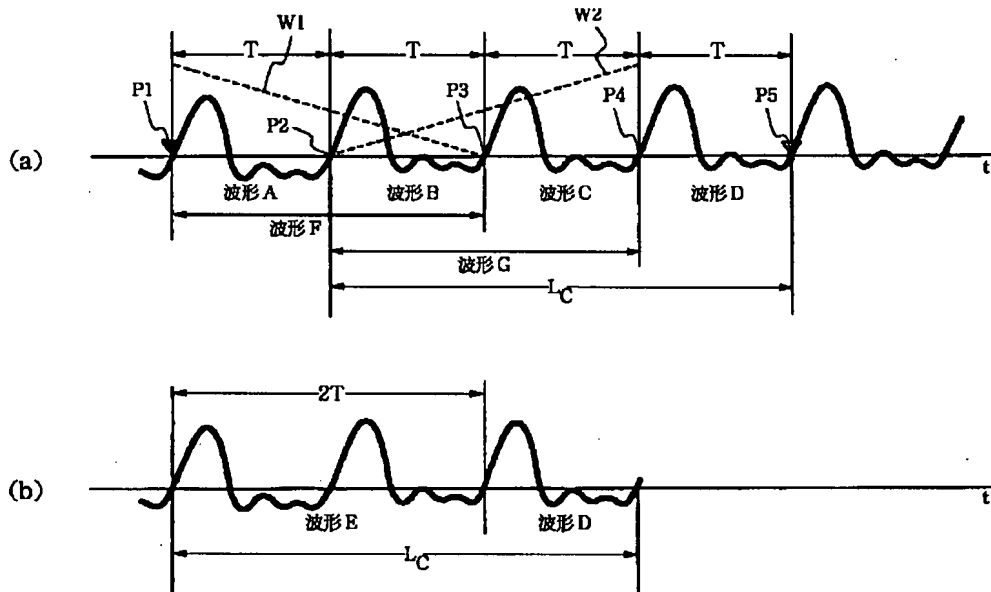
【図 1】



【図 4】



【図 2】



【図 3】

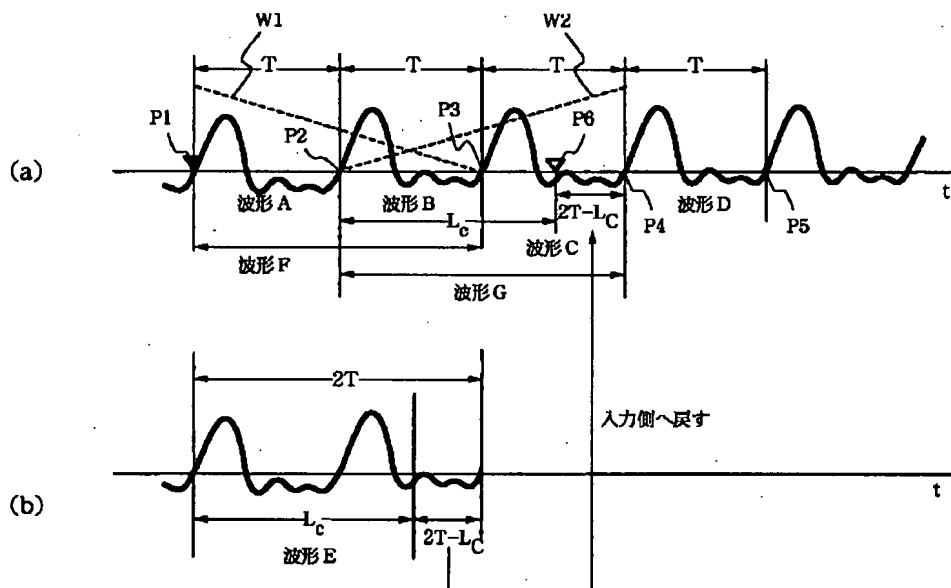


Figure 1 consists of two parts, (a) and (b), illustrating waveforms and time intervals.

Part (a) shows a sequence of waveforms labeled A, B, and C, separated by guard bands W3 and W4. Pilot tones P1 and P2 are indicated. The time interval between the start of waveform A and the start of waveform B is T. The time interval between the start of waveform B and the start of waveform C is T. The time interval between the start of waveform C and the end of the sequence is T. The total time interval between the start of waveform A and the end of the sequence is L_S . The time interval between the start of waveform A and the start of waveform B is $L_S - T$. The time interval between the start of waveform B and the start of waveform C is $L_S - T$. The time interval between the start of waveform C and the end of the sequence is $L_S - T$.

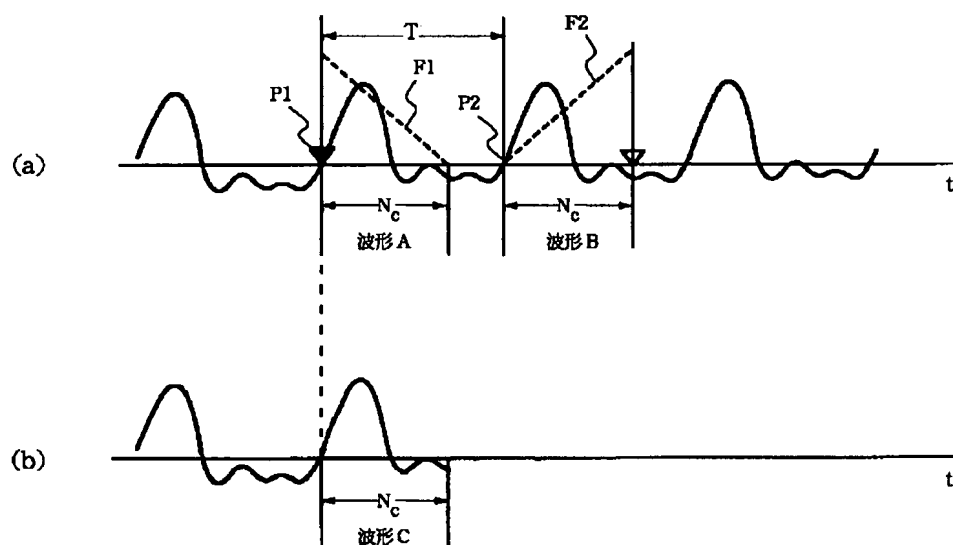
Part (b) shows a zoomed-in view of waveforms A, H, and C. The time interval between the start of waveform A and the start of waveform H is $2T$. The time interval between the start of waveform H and the start of waveform C is $L_S T$. The total time interval between the start of waveform A and the end of the sequence is $T + L_S$.

Figure 1 consists of two waveform diagrams, (a) and (b), illustrating the relationship between waveforms A, B, C, and H in a 3-phase system.

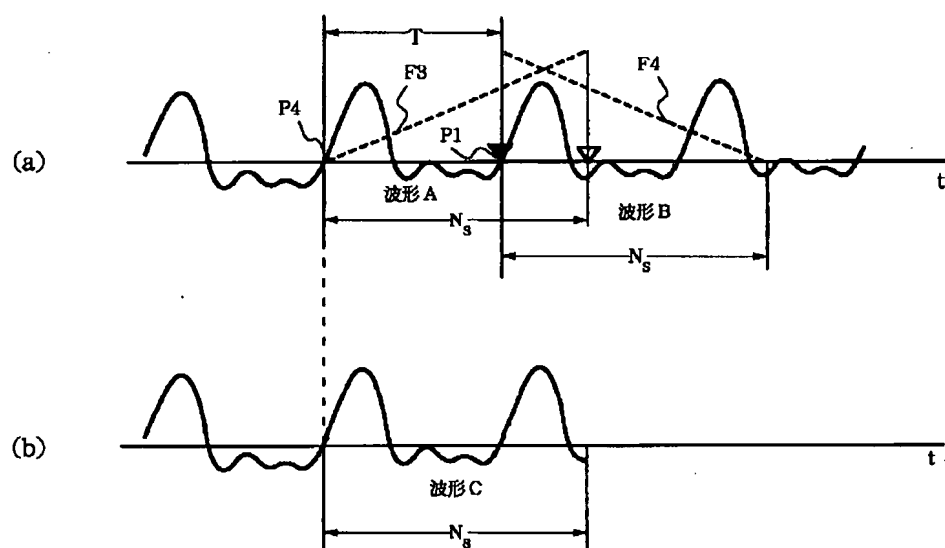
(a) Waveform A, B, and C: This diagram shows three waveforms labeled A, B, and C. The horizontal axis is time t . The period of the waveforms is T . The time interval between the start of waveform A and waveform B is L_s . The time interval between the start of waveform B and waveform C is $T - L_s$. The waveforms are labeled P4, W1, P1, P5, and W2. The waveforms are labeled 波形 A, 波形 B, and 波形 C.

(b) Waveform H: This diagram shows waveform H, which is a combination of waveforms A, B, and C. The horizontal axis is time t . The period of waveform H is $2T$. The time interval between the start of waveform A and waveform H is $T + L_s$. The time interval between the start of waveform H and waveform B is $T - L_s$. The waveforms are labeled 波形 A and 波形 H.

【図 7】



【図 8】



【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成10年(1998)11月4日

【公開番号】特開平7-181998

【公開日】平成7年(1995)7月21日

【年通号数】公開特許公報7-1820

【出願番号】特願平5-327898

【国際特許分類第6版】

G10L 9/14

【F I】

G10L 9/14 G
J

【手続補正書】

【提出日】平成9年4月22日

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0014

【補正方法】変更

【補正内容】

【0014】

【発明が解決しようとする課題】本発明は、上記のようにT D H Sを時間軸の圧縮あるいは伸長に適用すると圧縮率（あるいは伸長率） r が1に近い場合、 N_c や N_s がTより大きくなりすぎ加え合わせる音声波形内でピッチ周期の揺らぎによる波形のずれが生じ、歪みの原因となるという従来の問題点に鑑み、圧縮率（あるいは伸長率）が1に近い時でもずれや歪みの小さい音声波形が得られる音声圧縮及び伸長方式を開発することを目的とする。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0015

【補正方法】変更

【補正内容】

【0015】本発明は、音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して1ピッチ周期だけ進んだ位置との両方から夫々前記ピッチ周期の2倍以上且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記2つの音声信号波形を圧縮した1つの信号波形を生成し、次に前記現在ポイントを圧縮率に応じて前記入力された音声信号波形上で移動せしめて同様の操作を行い圧縮された音声信号を得るものである。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0016

【補正方法】変更

【補正内容】

【0016】また本発明は、音声信号の伸長時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して1ピッチ周期だけ遅れた位置との両方から夫々前記ピッチ周期の2倍以上で且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記2つの音声信号波形を合成した1つの信号波形を生成し、次に前記現在ポイントを圧縮率に応じて前記入力された音声信号波形上で移動せしめて同様の操作を行い伸長された音声信号を得るものである。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0020

【補正方法】変更

【補正内容】

【0020】そして、圧縮された音声信号 $s_c(n)$ をD/A変換器4に送り、ここでアナログ音声信号 $s_c'(t)$ に変換して出力する。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0030

【補正方法】変更

【補正内容】

【0030】そして、伸長された音声信号 $s_e(n)$ をD/A変換器4に送り、ここでアナログ音声信号 $s'(t)$ に変換して出力する。

【手続補正6】

【補正対象書類名】図面

【補正対象項目名】図4

【補正方法】変更

【補正内容】

【図4】

